SEQUENCE LISTING

<110> Schellenberger, Volker
Liu, Amy D.
Selifonova, Olga V.

<120> Directed Evolution of Microorganisms

<130> GC560-D1

<140> US 10/037,677
<141> 2001-10-23

<150> US 09/314,847
<151> 1999-05-19

<160> 17

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 741
<212> DNA
<213> Escherichia coli

<400> 1
atgaccgcta tgagcactgc aattacacgc cagatcgttc tcgataccga aaccaccggt 60
atgaaccaga ttggtgcgca ctatgaaggc cacaagatca ttgagattgg tgccgttgaa 120
gtggtgaacc gtcgcctgac gggcaataac ttccatgttt atctcaaacc cgatcggctg 180
gtggatccgg aagcctttgg cgtacatggt attgccgatg aatttttgct cgataagccc 240
acgtttgccg aagtagccga tgagttcatg gactatattc gcggcgcgga gttggtgatc 300
cataacgcag cgttcgatat cggctttatg gactacgagt tttcgttgct taagcgcgat 360
attccgaaga ccaatacttt ctgtaaggtc accgatagcc ttgcggtggc gaggaaaatg 420
tttcccggta agcgcaacag cctcgatgcg ttatgtgctc gctacgaaat agataacagt 480
aaacgaacgc tgcacggggc attactcgat gcccagatcc ttgcggaagt ttatctggcg 540
atgaccggtg gtcaaacgtc gatggctttt gcgatggaag gagagacaca acagcaacaa 600
ggtgaagcaa caattcagcg cattgtacgt caggcaagta agttacgcgt tgtttttgcg 660
acagatgaag agattgcagc tcatgaagcc cgtctcgatc tggtgcagaa gaaaggcgga 720
agttgcctct ggcgagcata a 741

<210> 2
<211> 246
<212> PRT
<213> Escherichia coli

<400> 2
Met Thr Ala Met Ser Thr Ala Ile Thr Arg Gln Ile Val Leu Asp Thr
1 5 10 15
Glu Thr Thr Gly Met Asn Gln Ile Gly Ala His Tyr Glu Gly His Lys
20 25 30
Ile Ile Glu Ile Gly Ala Val Glu Val Val Asn Arg Arg Leu Thr Gly
35 40 45
Asn Asn Phe His Val Tyr Leu Lys Pro Asp Arg Leu Val Asp Pro Glu
50 55 60
Ala Phe Gly Val His Gly Ile Ala Asp Glu Phe Leu Leu Asp Lys Pro

```
65                  70                  75                  80
Thr Phe Ala Glu Val Ala Asp Glu Phe Met Asp Tyr Ile Arg Gly Ala
                85                  90                  95
Glu Leu Val Ile His Asn Ala Ala Phe Asp Ile Gly Phe Met Asp Tyr
            100                 105                 110
Glu Phe Ser Leu Leu Lys Arg Asp Ile Pro Lys Thr Asn Thr Phe Cys
        115                 120                 125
Lys Val Thr Asp Ser Leu Ala Val Ala Arg Lys Met Phe Pro Gly Lys
    130                 135                 140
Arg Asn Ser Leu Asp Ala Leu Cys Ala Arg Tyr Glu Ile Asp Asn Ser
145                 150                 155                 160
Lys Arg Thr Leu His Gly Ala Leu Leu Asp Ala Gln Ile Leu Ala Glu
                165                 170                 175
Val Tyr Leu Ala Met Thr Gly Gly Gln Thr Ser Met Ala Phe Ala Met
            180                 185                 190
Glu Gly Glu Thr Gln Gln Gln Gln Gly Glu Ala Thr Ile Gln Arg Ile
        195                 200                 205
Val Arg Gln Ala Ser Lys Leu Arg Val Val Phe Ala Thr Asp Glu Glu
    210                 215                 220
Ile Ala Ala His Glu Ala Arg Leu Asp Leu Val Gln Lys Lys Gly Gly
225                 230                 235                 240
Ser Cys Leu Trp Arg Ala
                245
```

<210> 3
<211> 1164
<212> DNA
<213> Escherichia blattae

<400> 3

```
atgagctatc gtatgtttga ttatctggtt ccaaatgtga acttctttgg cccgggcgcc      60
gtttctgttg ttggccagcg ctgccagctg ctgggggtta aaaaagccct gctggtgacc     120
gataagggcc tgcgcgccat taaagacggt gctgtcgatc agaccgtgaa gcacctgaaa     180
gccgccggta ttgaggtggt cattttcgac ggggtcgagc cgaacccgaa agacaccaac     240
gtgctcgacg gcctggccat gttccgtaaa gagcagtgcg acatgataat caccgtcggc     300
ggcggcagcc cgcacgactg cggtaaaggc attggtattg cggccaccca cccgggtgat     360
ctgtacagct atgccggtat cgaaacactc accaacccgc tgccgcccat tattgcggtc     420
aacaccaccg ccgggaccgc cagcgaagtc acccgccact gcgtgctgac taacaccaaa     480
accaaagtaa aatttgtgat tgtcagctgg cgcaacctgc cttccgtctc cattaacgat     540
ccgctgctga tgatcggcaa gcccgccggg ctgaccgccg ccaccggtat ggatgccctg     600
acccacgcgg tagaggccta tatctccaaa gacgccaacc cggttaccga tgcctctgct     660
attcaggcca tcaaactgat tgccaccaac ttgcgccagg ccgtcgccct ggggaccaac     720
ctcaaagccc gtgaaaacat ggcctgcgcc tctctgctgg ccgggatggc ctttaacaac     780
gccaacctgg gctatgttca cgccatggct caccagctgg gcggcctgta cgacatggcc     840
cacggggtgg cgaacgcggt cctgctgccc catgtctgcc gctataacct gattgccaac     900
ccggaaaaat ttgccgatat cgccaccttt atgggggaaa acaccaccgg tctttccacc     960
atggacgcag cggagctggc catcagcgcc attgcccgtc tgtctaaaga tgtcgggatc    1020
ccgcagcacc tgcgtgaact gggggtaaaa gaggccgact tcccgtacat ggcagaaatg    1080
gccctgaaag acggcaacgc cttctctaac ccgcgcaaag ggaacgaaaa agagattgcc    1140
gacattttcc gccaggcatt ctga                                           1164
```

<210> 4
<211> 387
<212> PRT
<213> Escherichia blattae

<400> 4

```
Met Ser Tyr Arg Met Phe Asp Tyr Leu Val Pro Asn Val Asn Phe Phe
 1               5                  10                  15
Gly Pro Gly Ala Val Ser Val Val Gly Gln Arg Cys Gln Leu Leu Gly
            20                  25                  30
Gly Lys Lys Ala Leu Leu Val Thr Asp Lys Gly Leu Arg Ala Ile Lys
        35                  40                  45
Asp Gly Ala Val Asp Gln Thr Val Lys His Leu Lys Ala Ala Gly Ile
    50                  55                  60
Glu Val Val Ile Phe Asp Gly Val Glu Pro Asn Pro Lys Asp Thr Asn
65                  70                  75                  80
Val Leu Asp Gly Leu Ala Met Phe Arg Lys Glu Gln Cys Asp Met Ile
                85                  90                  95
Ile Thr Val Gly Gly Gly Ser Pro His Asp Cys Gly Lys Gly Ile Gly
            100                 105                 110
Ile Ala Ala Thr His Pro Gly Asp Leu Tyr Ser Tyr Ala Gly Ile Glu
        115                 120                 125
Thr Leu Thr Asn Pro Leu Pro Pro Ile Ile Ala Val Asn Thr Thr Ala
    130                 135                 140
Gly Thr Ala Ser Glu Val Thr Arg His Cys Val Leu Thr Asn Thr Lys
145                 150                 155                 160
Thr Lys Val Lys Phe Val Ile Val Ser Trp Arg Asn Leu Pro Ser Val
                165                 170                 175
Ser Ile Asn Asp Pro Leu Leu Met Ile Gly Lys Pro Ala Gly Leu Thr
            180                 185                 190
Ala Ala Thr Gly Met Asp Ala Leu Thr His Ala Val Glu Ala Tyr Ile
        195                 200                 205
Ser Lys Asp Ala Asn Pro Val Thr Asp Ala Ser Ala Ile Gln Ala Ile
    210                 215                 220
Lys Leu Ile Ala Thr Asn Leu Arg Gln Ala Val Ala Leu Gly Thr Asn
225                 230                 235                 240
Leu Lys Ala Arg Glu Asn Met Ala Cys Ala Ser Leu Leu Ala Gly Met
                245                 250                 255
Ala Phe Asn Asn Ala Asn Leu Gly Tyr Val His Ala Met Ala His Gln
            260                 265                 270
Leu Gly Gly Leu Tyr Asp Met Ala His Gly Val Ala Asn Ala Val Leu
        275                 280                 285
Leu Pro His Val Cys Arg Tyr Asn Leu Ile Ala Asn Pro Glu Lys Phe
    290                 295                 300
Ala Asp Ile Ala Thr Phe Met Gly Glu Asn Thr Thr Gly Leu Ser Thr
305                 310                 315                 320
Met Asp Ala Ala Glu Leu Ala Ile Ser Ala Ile Ala Arg Leu Ser Lys
                325                 330                 335
Asp Val Gly Ile Pro Gln His Leu Arg Glu Leu Gly Val Lys Glu Ala
            340                 345                 350
Asp Phe Pro Tyr Met Ala Glu Met Ala Leu Lys Asp Gly Asn Ala Phe
        355                 360                 365
Ser Asn Pro Arg Lys Gly Asn Glu Lys Glu Ile Ala Asp Ile Phe Arg
    370                 375                 380
Gln Ala Phe
385
```

<210> 5
<211> 1164
<212> DNA
<213> Escherichia blattae

<400> 5

```
atgagctatc gtatgtttga ttatctggtt ccaaatgtra acttctttgg cccgggcgcc      60
gtttctgttg ttggccagcg ctgccagctg ctggggggta aaaaagccct gctggtgacc     120
gataagggcc tgcgcgccat taaagacggt gctgtcgatc agaccgtgaa gcacctgaaa     180
gccgccggta ttgaggtggt cattttcgac ggggtcgagc cgaacccgaa agacaccaac     240
gtgctcgacg gcctggccat gttccgtaaa gagcagtgcg acatgataat caccgtcggc     300
ggcggcagcc cgctcgactg cggtaaaggc attggtattg cggccaccca cccgggtgat     360
ctgtacagct atgccggtat cgaaacactc accaacccgc tgccgcccat tattgcggtc     420
aacaccaccg ccgggaccgc cagcgaagtc acccgccact gcgtgctgac taacaccaaa     480
accaaagtaa aatttgtgat tgtcagctgg cgcaacctgc cttccgtctc cattaacgat     540
ccgctgctga tgatcggcaa gcccgccggg ctgaccgccg ccaccggtat ggatgccctg     600
acccacgcgg tagaggccta tatctccaaa gacgccaacc cggttaccga tgcctctgct     660
attcaggcca tcaaactgat tgccaccaac ttgcgccagg ccgtcgccct ggggaccaac     720
ctcaaagccc gtgaaaacat ggcctgcgcc tctctgctgg ccgggatggc ctttaacaac     780
gccaacctgg gctatgttca cgccatggct caccagctgg gcggcctgta cgacatggcc     840
cacggggtgg cgaacgcggt cctgctgccc catgtctgcc gctataacct gattgccaac     900
ccggaaaaat ttgccgatat cgccaccttt atgggggaaa acaccaccgg tctttccacc     960
atggacgcag cggagctggc catcagcgcc attgcccgtc tgtctaaaga tgtcgggatc    1020
ccgcagcacc tgcgtgaact gggggtaaaa gaggccgact tcccgtacat ggcagaaatg    1080
gccctgaaag acggcaacgc cttctctaac ccgcgcaaag ggaacgaaaa agagattgcc    1140
gacattttcc gccaggcatt ctga                                           1164
```

<210> 6
<211> 387
<212> PRT
<213> Escherichia blattae

<400> 6

```
Met Ser Tyr Arg Met Phe Asp Tyr Leu Val Pro Asn Val Asn Phe Phe
 1               5                  10                  15
Gly Pro Gly Ala Val Ser Val Val Gly Gln Arg Cys Gln Leu Leu Gly
            20                  25                  30
Gly Lys Lys Ala Leu Leu Val Thr Asp Lys Gly Leu Arg Ala Ile Lys
        35                  40                  45
Asp Gly Ala Val Asp Gln Thr Val Lys His Leu Lys Ala Ala Gly Ile
    50                  55                  60
Glu Val Val Ile Phe Asp Gly Val Glu Pro Asn Pro Lys Asp Thr Asn
65                  70                  75                  80
Val Leu Asp Gly Leu Ala Met Phe Arg Lys Glu Gln Cys Asp Met Ile
                85                  90                  95
Ile Thr Val Gly Gly Gly Ser Pro Leu Asp Cys Gly Lys Gly Ile Gly
            100                 105                 110
Ile Ala Ala Thr His Pro Gly Asp Leu Tyr Ser Tyr Ala Gly Ile Glu
        115                 120                 125
Thr Leu Thr Asn Pro Leu Pro Pro Ile Ile Ala Val Asn Thr Thr Ala
    130                 135                 140
Gly Thr Ala Ser Glu Val Thr Arg His Cys Val Leu Thr Asn Thr Lys
145                 150                 155                 160
Thr Lys Val Lys Phe Val Ile Val Ser Trp Arg Asn Leu Pro Ser Val
                165                 170                 175
Ser Ile Asn Asp Pro Leu Leu Met Ile Gly Lys Pro Ala Gly Leu Thr
            180                 185                 190
Ala Ala Thr Gly Met Asp Ala Leu Thr His Ala Val Glu Ala Tyr Ile
        195                 200                 205
Ser Lys Asp Ala Asn Pro Val Thr Asp Ala Ser Ala Ile Gln Ala Ile
    210                 215                 220
Lys Leu Ile Ala Thr Asn Leu Arg Gln Ala Val Ala Leu Gly Thr Asn
225                 230                 235                 240
```

```
Leu Lys Ala Arg Glu Asn Met Ala Cys Ala Ser Leu Leu Ala Gly Met
                245                 250                 255
Ala Phe Asn Asn Ala Asn Leu Gly Tyr Val His Ala Met Ala His Gln
            260                 265                 270
Leu Gly Gly Leu Tyr Asp Met Ala His Gly Val Ala Asn Ala Val Leu
        275                 280                 285
Leu Pro His Val Cys Arg Tyr Asn Leu Ile Ala Asn Pro Glu Lys Phe
    290                 295                 300
Ala Asp Ile Ala Thr Phe Met Gly Glu Asn Thr Thr Gly Leu Ser Thr
305                 310                 315                 320
Met Asp Ala Ala Glu Leu Ala Ile Ser Ala Ile Ala Arg Leu Ser Lys
                325                 330                 335
Asp Val Gly Ile Pro Gln His Leu Arg Glu Leu Gly Val Lys Glu Ala
            340                 345                 350
Asp Phe Pro Tyr Met Ala Glu Met Ala Leu Lys Asp Gly Asn Ala Phe
        355                 360                 365
Ser Asn Pro Arg Lys Gly Asn Glu Lys Glu Ile Ala Asp Ile Phe Arg
    370                 375                 380
Gln Ala Phe
385
```

<210> 7
<211> 12
<212> DNA
<213> Artificial Sequence

<220>
<223> wild type mutD gene

<400> 7
atgaccgcta tg 12

<210> 8
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> pOS100 mutD mutated gene

<400> 8
ttgacgcttt g 11

<210> 9
<211> 12
<212> DNA
<213> Artificial Sequence

<220>
<223> pOS101 mutD mutated gene

<400> 9
gtgaccgctg tg 12

<210> 10
<211> 11
<212> DNA

<213> Artificial Sequence

<220>
<223> pOS102 mutD mutated gene

<400> 10
gtgccgctgt g 11

<210> 11
<211> 12
<212> DNA
<213> Artificial Sequence

<220>
<223> pOS104 mutD mutated gene

<400> 11
ttgaccgctt tg 12

<210> 12
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<223> pOS105 mutD mutated gene

<400> 12
gtgaccgctg tgagcacttg caattacacg ccagatcgtt ctcgataccg aaatc 55

<210> 13
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> pOS106 mutD mutated gene

<400> 13
gtgaccgctt g 11

<210> 14
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 14
cgcctccagc gcgacaatag cggccatc 28

<210> 15
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 15
ccgactgaac taccgctccg cgttgtg 27

<210> 16
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 16
tctgatacgg gatcctcaga atgcctggcg gaaaat 36

<210> 17
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 17
gcgccgtcta gaattatgag ctatcgtatg tttgattatc tg 42